redundant or ineffective features may be discarded. Since subjectively rated video imagery was unavailable at the time of writing this report, emphasis has been placed on development of a candidate set of features for automated quality assessment of digitally transmitted video.

2. DESCRIPTION OF FEATURES

The most difficult process in virtually all pattern recognition and classification systems is feature extraction. A general theory of feature extraction is unavailable and most feature extraction methods are ad hoc and highly application dependent. The performance of a classifier is determined primarily by the features that are injected into the classifier. For this reason, the bulk of the development work for a classification system is to develop methods that extract sensitive and relevant feature values. This section describes the development of a set of features for automatically assessing the quality of digitally transmitted video. Emphasis has been placed on automated techniques for cost effective monitoring, and repeatability.

To understand the features that have been developed, background information is first presented on common video artifacts, desirable properties of features, and proper alignment of original and distorted video imagery. Techniques for video scene alignment, very rarely covered in the literature, are discussed in section 2.3. Calculation of some features requires proper temporal alignment of original and distorted video imagery.

Rationale for preconditioning the sampled video before feature extraction is discussed. The technique for extracting each feature from the sampled video is described in detail. The features objectively quantify the presence of common video artifacts. Of critical concern here is the computational time of a particular feature. Alternate algorithms are presented that reduce this cost of computation. For illustrative purposes, each feature extraction technique is demonstrated using VTC/VT data.

2.1 Common Video Compression Artifacts

The American National Standards Institute, Accredited Standards Committee T1, Working Group T1Q1.5 is drafting interface performance specifications for digital VTC/VT and digital television. The VTC/VT

sub-working group of T1Q1.5 is developing a catalogue of video motion artifacts associated with video compression and the resultant effects on video quality. The motion artifacts that are most noticeable to the viewer and that show the most potential for being measured are reproduced in Table 1. The artifact, definition of the artifact, and examples of the artifact are listed in the table. Artifacts are most apparent when video motion is present. The information content of a video signal that contains moving and/or changing scenes may simply be too great for a fixed transmission data rate. In such cases, image pixel values may not be updated rapidly enough, resulting in noticeable artifacts. Additional video coding artifacts can be found in Murakami et al. (1988).

Probably the most noticeable and objectional motion artifact is resolution degradation. Normally, stationary objects are coded with relatively high spatial resolution. However, as soon as the object moves, blurring and/or jerky motion of the object is noticed. In cases of excessive motion such as during camera pans and zooms, very objectionable blocking artifacts may appear. Other image coding artifacts seen upon close inspection include edge busyness and image persistence.

Table 1. Common Video Compression Artifacts

Motion Artifact	**Definition**
1. Resolution Degradation Examples:	The deterioration of motion video such that the received video imagery has suffered a loss of spatio-temporal resolution.
Blocking	The received video imagery possesses rectangular or checkerboard patterns not present in the original.
Blurring/smearing	The received video imagery has lost edges and detail present in the original.
Jerkiness	The original smooth and continuous motion is perceived as a series of distinct snapshots.
2. Edge Busyness Example:	The deterioration of motion video such that the outlines of moving objects are displayed with randomly varying activity.
Mosquito noise	The quantizing noise generated by the block processing of moving objects that gives the appearance of false small moving objects (e.g. a mosquito flying around a person's head and shoulders).
3. Image persistence Example:	The appearance of earlier faded video frames of a moving and/or changing object within the current video frame.
Erasure	An object that was erased continues to appear in the received video imagery.

2.2 Desirable Properties of Features

For the video quality measurement system shown in Figure 1, developing a set of sensitive and relevant features can be very difficult. Often, intuition and ad hoc procedures must be used to obtain a set of features which are meaningful and easily computed. The following list details some desirable properties of objectively measured features. These properties were used to steer the development of a set of features for measuring the quality of digitally transmitted video.

1. Correlation with subjective quality

   Perhaps the most critical attribute of a meaningful feature is strong correlation of the measured feature value with the subjective rating. If overall subjective ratings are not available, features should at least be sensitive to the amount of subjectively noticed video artifacts. The feature value should change monotonically when the amount of the artifact or distortion is increased.

2. Automation

   Feature extraction should be performable by an autonomous measurement system. Advantages include automatic detection of transmission line impairments, cost effective monitoring, and repeatability.

3. Application to many types of scenes

   Since the performance of the digital compression and transmission algorithm normally depends upon the type of imagery which is being compressed, the feature extraction procedure should be applicable to arbitrary video scenes. Thus, to test the video quality performance for a specific user application, one must use the appropriate type of video scenes.

4. Application as a local estimate

   There is evidence that the human viewer may determine the quality of a video scene by rating the quality of local details within the video scene (Westernik and Roufs, 1988).

Thus, the human viewer will often look at high contrast edges and contours to perform quality judgments. To account for this phenomena, feature extraction methods should take into account local or sub-regional properties (in space and/or time) of the video. Local estimates of quality may also be utilized by video compression algorithms to allocate bits dynamically to each sub-region of the video image.

5. Computational efficiency

Features that are rapidly computed from the image are preferable from a cost and implementation standpoint. At best, the feature should be computable in real time, given reasonable hardware. Computationally efficient features may also be required for large, higher resolution imagery, such as HDTV.

6. Stability

The feature should not be sensitive to distortions which the human viewer does not notice. For example, the feature should not be sensitive to small shifts in the mean of the video imagery nor other image distortions which fall below the threshold of visibility.

7. Functional independence

When choosing a feature set, every feature within the set should convey different information. If a particular feature can be obtained as a function of other features within the feature set, that feature does not convey any additional information and can be disregarded.

8. Technology independence

The feature is useful for a wide range of technologies. For instance, a feature developed for measuring digital image compression artifacts should also be useful in measuring video quality of an analog transmission channel.